DOCUMENT RESUME

ED 082 455 EM 011 384

AUTHOR	Merrill, M. David; Boutwell, Richard C.
TITLE	Instructional Development: Methodology and Research. Working Paper Number 33.
INSTITUTION	Brigham Young Univ., Provo, Utah. Dept. of Instructional Research and Development.
PUB DATE	31 May 72
NOTE	76p.; Reprint from The Review of Research in Education, Fred N. Kerlinger; See Also EM 011 385
AVAILABLE FROM	Brigham Young University Press, 209 University Press Building, Provo, Utah 84601 ($2.50)
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	*Curriculum Development; *Educational Research; *Higher Education; *Instructional Design; Learning Theories; Methodology; Publications; *Research Methodology; Task Analysis
IDENTIFIERS	Mathemagenic Information

ABSTRACT

Five premises necessary for the formulation of instructional theory are stated. As a suggested elaboration of the "behavior classification" premise, a two-way (content and behavior) classification scheme for higher order tasks is described. The content dimension includes paired associate, concept, principle, and problem tasks. The behavior dimension includes discriminated recall, classification, rule using, and higher rule using. Each two-way category is illustrated by treatments from recent research on higher order tasks. As a suggested elaboration of the "manipulation of task variables" premise, a taxonomy of task variables is described. The taxonomy is divided into qualitative variables and quantitative parameters. The former include the major categories of presentation form, interdisplay relationships, and mathemagenic information (i.e. information designed to promote behavior which gives birth to learning). Parameters include major categories of sequence, quantity, and pace. Research investigating some of the variables in the taxonomy is described and critiqued, and some additional needed research is suggested. A final section indicates implications of this research and the task taxonomy for instructional development. (Author)

ED 082455

EM 011 384

IRD

Department of Instructional Research and Development
Division of Instructional Services - Brigham Young University

INSTRUCTIONAL RESEARCH AND DEVELOPMENT

INSTRUCTIONAL DEVELOPMENT: METHODOLOGY AND RESEARCH

M. David Merrill
and
Richard C. Boutwell

Working Paper No. 33
May 31, 1972

This is a prepublication draft designed for limited circulation. The reader's comments are solicited. Please do not cite as a source without consulting the department and the author of this document.

Division of Communication Services
Brigham Young University
Provo, Utah

INSTRUCTIONAL DEVELOPMENT: METHODOLOGY AND RESEARCH

M. David Merrill
Brigham Young University

Richard C. Boutwell
Brigham Young University

Within the past few years, a major portion of the money available for educational improvement has been channeled into the systematic development of instructional systems and products. Several books and manuals have appeared giving guidance to potential developers of instructional systems. (Briggs, 1970; Popham & Baker, 1970; Gerlack & Ely, 1971; Kemp, 1971; Merrill, 1971a). While each varies in terminology and detail, every position includes the following components: behavioral objectives, pretest, instructional activities, posttest, and revision based on empirical tryouts. This model was first given widespread exposure by Robert Glaser in his papers describing the emerging field of instructional psychology (Glaser 1965, 1966).

Merrill (1972a, 1972b) suggested that the following premises underlie this basic model:

<u>Premise 1</u>: Objectives must be specified in terms of observable student behavior.

<u>Premise 2</u>: Testing instruments should measure the student's ability to perform specified behavior (criterion referenced) rather than how well he performs in comparison with other students (norm referenced). (See Glaser 1963; Glaser & Klaus, 1962; and Glaser & Nitko, 1971)

<u>Premise 3</u>: Instructional products must be verified by empirical procedures.

The paucity of research related to these basic premises suggests that they are considered axiomatic by most instructional developers. Numerous propositions have been suggested for each of these statements, but there are only a handful of rather poorly executed studies testing them. Space restrictions prevent a detailed review of this area (e.g. studies related to behavioral objectives include the following: Mager & McCann, 1961; McNeil, 1967; Baker, 1969; Horton, 1969; Cook, 1970; Oswald, 1970; Dalis, 1970; Jenkins & Deno, 1971; Jenkins & Neisworth, 1971; Paul Merrill, 1971; Hiscox & Herron, 1971; Pratt, 1971; Rosen, 1971).

Most instructional development efforts can be characterized by a "raw empiricism" approach. Instructional materials are prepared based on intuition, folklore, and experience. These materials are then administered to members of the target population. If the students pass the test, the product is considered appropriate; if not, the materials are revised and tried again. This tryout-revision cycle is repeated until the product works or the developers run out of resources or time.

Progress cannot be made toward an empirically based instructional development methodology until propositions that relate objectives to learning activities are specified and tested. In order to state such propositions, it is necessary to propose at least two additional premises:

Premise 4: Instructional outcomes can be classified into a limited number of behavioral classes.

Premise 5: The acquisition of a given kind of behavioral outcome can be optimized by the appropriate manipulation of task variables.

A series of previous reviews have surveyed the research literature related to instructional psychology (Anderson, 1967; Gagné & Rohwer, 1969; and Glaser & Resnick, 1972). While most of the instruction relevant to school learning involves classification, rule using, or problem solving, much of the previously reviewed research deals with recall. Frequently, investigators have not made a distinction between various kinds of instructional outcomes and have tended to investigate simple tasks and then generalize their findings to all learning situations. The purpose of this paper is to provide an interpretative essay and research review concentrating on tasks which require these more complex behaviors.

A difficult methodological task facing any researcher is the adequate definition of his independent and dependent variables. This problem is confounded when the variables are complicated and when researchers use as variables involved procedures developed by the practitioner (e.g., discovery and inquiry). The first section of this paper will propose a two-dimensional classification system for determining a behavioral and content category. This system follows from Premise 4 above, and provides a framework for interpreting existing research and a method for more adequately defining independent variables for future research.

The second section proposes a taxonomy of task variables. Perhaps the most salient factor contributing to the paucity of research in this area is the lack of any systematic identification of those variables that are manipulated by instructional developers. The proposed taxonomy makes it possible to compare research studies related to Premise 5, and provides a method for defining variables in future research so that meaningful prescriptive propositions can be stated and tested.

The elements of both of the previously defined schemes are illustrated with tasks and treatments used by researchers in this area. The third section describes a cross section of this research using the categories thus defined and suggests many propositions that have not yet been investigated. A final section suggests implications of these schemes and resulting propositions for instructional development.

Two-Dimensional Task Classification

Existing schemes for classifying cognitive behavior differ in that each categorizes a different aspect of the instruction-learning task. Bloom and his co-workers (Bloom, Englehart, Faust, Hill & Krahwohl, 1956) based their categories on a combination of content and behavior, indicating that the system was a taxonomy of educational objectives. Gagné (1965, 1970) separated categories on the basis of the conditions required to promote a given type of behavior. Merrill (1971a, 1971b) modified and extended the Gagné categories, but used as a basis for classification intended overt behavior and the conditions under which it must be observed.

A more complete taxonomy can be described if one recognizes that there are two somewhat independent phenomena that can be classified. The first, task content, refers to those characteristics of a task identifying it as primarily a paired associate, concept, principle (rule), or problem task. The second, student behavior, refers to the overt acts a student performs and the conditions under which these acts must be observed before they can be termed discriminated recall, classification behavior, rule-using behavior, or higher-order rule-using behavior.

The content dimension is independent of the behavior dimension in that a task may be classified at one content level (e.g., concept), but the student can still behave in relation to this task using any of several behavior levels (e.g., discriminated recall or classification). Content and behavior are not completely independent, however, in that the content of a given task limits the level of behavior that can be demonstrated. For example, it is not possible to demonstrate classification behavior in relation to a paired associate task. However, it is possible to respond using a lower level of behavior than that most appropriate for the content.

Figure 1 illustrates this relationship of the content and behavior dimensions of task classification. The boxes on the diagonal of Figure 1 are the categories previously identified for the Gagné-Merrill hierarchy. The italicized words suggest new terminology that can refer to a task simultaneously classified on both dimensions. The other boxes indicate combinations of task content and behavior and suggest some of the possible terms that can refer to such combinations.

Insert Figure 1 About Here

Task Content

Paired Associate Content. The content of a task is said to be paired associate when a set of symbols, objects, or events is associated on a one-to-one basis with another set of symbols, objects, or events. For a one-to-one association, the association must be made, not with just any instance of the two sets, but with a particular member of one set paired with a particular member of another set.

Some everyday paired associate tasks include the following examples: letters of the alphabet with letter names (symbol-to-symbol); notes on a musical staff with note names (symbol-to-symbol); parts of a sewing machine with their names (object-to-symbol); inventions with inventors [objects-to-symbols (names) or objects-to-objects (pictures)], sports records with the record holder [events-to-symbols (names)].

The task can be organized in list form or in paragraph form and still retain the critical characteristics of a paired associate task. One of the recent areas of substantial research activity is the investigation of "prose learning" by Rothkopf, Frase, and others, which was reviewed by Glaser and Resnick (1972). The "prose" experimental materials used in much of this research consisted of paired associate tasks organized in paragraph form as in the following sample (Frase, 1968a);

> "Jim is a pilot. He was born in 1921.
> John is a policeman. He was born in 1930.
> Jack is a butcher. He was born in 1926.
> Jeff is an engineer. He was born in 1934."

Concept Content. The content of a task consists of a concept when it includes a concept definition and a set of discriminably different symbols, objects, or events, which all have in common one or more attributes. A concept definition is a list of relevant attributes. Attributes are those characteristics which determine class membership. Characteristics may be physical, functional, or relational. A concept is identified by a concept name--usually a symbol which refers to the members of the class individually or collectively.

Tennyson, Woolley, and Merrill (1972) used a real world concept of "trochaic meter." The set of instances includes all passages of poetry with a stressed syllable followed by an unstressed syllable. The relevant attributes are themselves concepts, there are an indefinitely large number of instances, and there is a large and varied number of irrelevant attributes that may or may not be present in every instance (e.g., number of lines, rhyme, subject, meaning, language, and punctuation).

Principle Content. Much of the confusion surrounding instruction and research on tasks above the concept level is caused by the lack of a satisfactory definition of principles. Gagné (1970) defined a principle as "... an inferred capability that enables the individual to respond to a class of stimulus situations with a class of performances, the latter being predictably related to the former by a class of relations." The formal definition of a rule proposed by Scandura (1966, 1968, 1970) is: "... a set of stimulus properties (D) which determine responses, a set of response properties (R), and an operation (O) between them, such that each element in the first set is associated with exactly one element in the second..." As Scandura points out, this is the definition of a mathematical function and has the following form: R = 0 (D) where D is the domain, 0 the function, and R the range.

Extending Scandura's notion, a principle task consists of a stated rule in such a form that the stimulus set (D-set) and the response set (R-set) are identified and the operation is specified. It also consists of a set of problems. A problem is defined as the presentation of a member of the D-set with instructions to the student to apply the operation thus producing one and only one member of the R-set.

Guthrie (1967) taught several rules for solving cryptograms. The following restatement of one of his rules illustrates the above definition of a principle task:

> Each $\boxed{\text{Type I}}$ cryptogram (D-set) can be changed into a meaningful word (R-set) by exchanging the first and the last letters (operation).

The problem set consists of all cryptograms.

Problem Content. A higher order problem is one that involves the use of two or more rules, one of which determines the use of the other. Scandura has suggested in his more recent formulations (1968, 1970) that it is possible to have a rule for determining which operation (rule) to use in a given situation. If a principle is a statement that relates a D-set to an R-set by means of an operation (0); then a higher order principle is a statement that indicates an operation for selecting one D-set from a set of D-sets, a corresponding R-set from a set of R-sets, and an appropriate operation from a set of operations.

A problem task consists of such a higher order statement and a set of problems that can be solved by application of this rule. The following example comes from Roughead and Scandura (1968). The set of D-sets consisted of all possible number series. The set of R-sets consisted of all possible sums of number series. The 0-set consisted of summing formulas for each type of number series. The higher order rule given is as follows:

> ...formulas (0 set) for $\underline{\sum n}$ (set of D-sets) may be written as a product of an expression involving $\underline{n}$ [i.e., f (n)] and $\underline{n}$ itself, (higher order 0).... The required formula is simply $\underline{\sum n = n \cdot f\ (n)}$. n = number of terms in a given series.

One of the specific rules that can be determined by application of the higher order rule is as follows:

> The sum (R-set) of a number series starting with 1 and consisting of consecutive odd numbers (D-set) can be determined by squaring the number of terms (0).

Student Behavior Related to Task Content

Discriminated Recall. A student is responding with discriminated recall behavior when, given a member of the stimulus set in any order--a symbol, object, or event--he can immediately respond by providing the associated symbol, or indicating the associated object or event. The word recall is used in the broad sense synonymous with memory. Recall is sometimes used to mean a constructed versus a recognized response. While there are demonstrated differences between recognizing and constructing a response, both behaviors are included in the discriminated recall category. The first critical condition for discriminated recall is that S must be given a member of a set which he has previously been shown during learning. Second, the response must occur with a very short latency (immediately). Third, the student must be able to respond to the stimulus terms presented in any order; otherwise he may be exhibiting serial recall rather than discriminated recall. In addition, discriminated recall can be directly observed; that is, no inference is necessary. Either a student remembers the associations within a limited latency period or he does not. Except to measure retention, it is only necessary for the student to make a single response to adequately demonstrate any given association in the set.

Discriminated Recall in Concept Principle and Problem Tasks. Embedded within every concept task are two paired associate tasks, definition and instance recall. First, the student can be asked to recognize or restate the definition or several definitions. Anderson (1972a) suggests that four types of test questions are possible in this situation: verbatim questions, statements taken word-for-word from the text; transformed verbatim questions, statements rearranged into various syntactical forms; paraphrase questions, statements in which synonyms are substituted in the original statements; and transformed paraphrase questions, where a paraphrased statement is rearranged into various syntactical forms. Anderson suggests that more meaningful processing is required by paraphrased questions. While probably requiring at least a two-step memory process, the authors would still classify paraphrased definition recall as discriminated recall.

The second discriminated recall behavior embedded in every concept task is the recall of previously encountered instances. If a student is presented a set of instances and told they are members of a particular class, and then, in a test situation, is asked to recognize these same class members, the behavior required is not classification, but discriminated instance recall. In practical situations, students are often asked to remember instances of a concept from their past experience or to bring examples from home. In most cases, the instances thus generated are those that were presented during instruction or that closely resemble those presented during instruction and that require a form of discriminated instance recall.

A student can be asked to demonstrate discriminated recall in a principle task as well as in a concept task. In addition to definition and instance recall indicated for concept tasks, he can be asked to recall or recognize the stated rule. Whether recognized, reproduced word-for-word, or paraphrased, remembering a stated rule is primarily discriminated recall.

If the operation has already been applied to a particular member of the D-set and this application shown to the student, then asking him to resolve the problem is a solution-recall situation, not a rule-using situation.

For higher order problem tasks, there are two more recall possibilities. S can be asked to recall the statement of the higher order rule or the application of the higher order rule to a particular D-set.

Definition and rule recall are the most frequently used testing procedures in most instructional settings even when classification or rule-using behavior would provide a better situation for inferring comprehension. Clearly identifying the separate content and behavior components in these situations should facilitate more adequate evaluation.

Classification Behavior. S is demonstrating classification behavior when, given an unencountered instance (exemplar or nonexemplar) of a particular class, he is able to indicate class membership. S can indicate class membership by giving the class name or pointing to the exemplar when given an exemplar-nonexemplar pair. The critical conditions under which classification behavior must be observed include the following: first, the instance must be unencountered; that is, it must not have been previously identified as an exemplar or nonexemplar of the class being taught. Second, it is not possible or practical to observe S's response to every possible instance since real world concepts consist of indefinitely large classes. Therefore, when his classification of one instance is observed, we must make inferences about his ability to correctly classify yet unencountered instances. In this sense, any observation of classification behavior must take place in a transfer situation. Third, since we are making inferences about S's ability to correctly classify yet unencountered instances, the adequacy of this inference depends on how well the sample instances represent the range of instances included in the class. Fourth, since we are making inferences from a sample, it is necessary for reliable observation to observe S's classification of several instances. Up to a point, the more instances used for observation, the more reliable the inferences.

S can also demonstrate his understanding of a particular class, if, during presentation of instances the relevant attributes are not named, but later he is asked to identify them after he has been shown a number of exemplars and nonexemplars. This technique of measurement is the most frequently used procedure for laboratory investigations of concept learning (see Clark, 1971), but has not been used widely in more practical situations.

Anderson's (1972a) fifth and sixth kind of questions for measuring comprehension deal directly with classification measurement. His fifth type is "questions formed by substituting particular for general terms." The name and definition of a concept are general terms while the instances are particular terms. This is equivalent to presenting new instances to the student and having him classify them. The final type of question suggested by Anderson consists of "questions formed by substituting general terms for specific terms." This is equivalent to deriving the definition from instances which have been presented.

Classification in Principle or Problem Tasks. Even though the content of the tasks is principles, S may still be required to demonstrate classification behavior rather than application of the operation. Classification behavior can be demonstrated by having S indicate class membership for instances of either the stimulus set (D-set) or response set (R-set).

When the content of a task is higher order problems, S may also be asked to classify an entire D-set or an entire R-set as members of the greater sets included as part of the higher order problem situation under consideration.

In addition, he can classify instances of the operation.

Principle or problem classification is frequently neglected in instructional evaluation. Students are often able to use a given operation once a problem has been identified as a member of the D-set to which the operation applies but they are unable to determine when to use a given operation. Identifying the classification components of principle tasks should facilitate more adequate evaluation.

Rule-Using Behavior. S is demonstrating rule-using behavior if, when given an unencountered instance of a particular D-set, he can apply the operation (0) and produce the corresponding member of the R-set. S can demonstrate this behavior by writing or pointing to the appropriate member of the R-set, or he can be given a particular member of the R-set and the operation and asked to which member of the D-set the operation should be applied. (There are, however, some noncommutative rules for which this procedure would be inappropriate.) Unless S is told that the particular problem presented is a member of the appropriate D-set, he must first perform a classification task by correctly identifying the class membership of the instance. The critical conditions under which rules must be observed for adequate inference are similar to concepts and require unencountered problems, inference from a transfer situation, and adequate problem sampling.

Anderson (1972b) gives a good example of rule-using behavior applied to a principle task, using the following principle: "Intermittent reinforcement causes high resistance to extinction." Ss were presented this principle in a paragraph that further defined "intermittent reinforcement" and "resistance to extinction," and that presented an example of the application of the principle. The test consisted of presenting a series of multiple-choice questions putting a particular organism in a given situation (D-set) and indicating a particular reinforcement schedule (0). The alternatives were various patterns of potential resulting response (R-set). An example is as follows:

> "A hungry pigeon in an experiment is given a pellet of food for some, but not all, pecks on an illuminated disc. When food is no longer given, the pigeon will:
>
> a. squawk and flap its wings.
> b. begin to peck the walls and floor of the cage.
> c. begin to peck the disc more frequently.
> d. maintain pecking the disc.
> e. soon stop pecking the disc."

The sample of items presented varied the organisms, situations, and the patterns of reinforcement. Compared to a control group who read a control passage, the experimental Ss performed better. The performance was best on items identical to the sample given in the passage (discriminated recall), next best on very similar items, and worse on different items (rule using).

As with classification behavior, it is possible to infer comprehension of a principle under an inductive situation where S is first presented a number of solved problem instances (instances that illustrate the application of 0 to particular members of the D-set) and then asked to indicate the appropriate operation. If during the process of presenting solved instances, S is told the rule, he may merely demonstrate discriminated recall. Even after he has discovered the rule following such a procedure, it is desirable to test him in unencountered problem situations.

Rule Using in Problem Tasks. When the content consists of higher-level problems, S can be asked to apply one of the operations constituting the O-set to its D-set. When the rule is identified for the student rather than asking him to generate the rule by application of the higher order rule, the resulting behavior is rule-using rather than higher order rule-using.

Higher Order Rule-Using Behavior. S is demonstrating higher order rule-using behavior when, given an unencountered D-set, he can apply the higher order rule and produce the operation to be applied to each member of that D-set to produce appropriate members of the R-set. To observe higher order rule-using behavior, S can be given an R-set and operation and asked to determine the appropriate D-set. (As with rules, this is not appropriate in all cases.) The D-set must be an unencountered set, and the application of the higher order rule must be to a new situation. Constraints of adequate sampling and reliability also apply.

Taxonomy of Task Variables

Research-based prescriptions for the manipulation of task variables are scant in spite of the central position Premise 5 has for most of the current work in instructional development. Perhaps one reason for this limited activity is that a systematic scheme for identifying which task variables are most likely to effect instruction has not previously been developed. This paper proposes such a classification scheme.

Table 1 summarizes the taxonomy. Three major qualitative categories are proposed--presentation form, interdisplay relationships, and mathemagenic information. Each of these categories is subdivided into several variables and each variable is further subdivided into two values representing ends of a set of qualitative continua. Three quantitative parameters are also proposed --sequence, quantity, and pace. Each parameter consists of several quantitive dimensions which can be applied to any of the qualitative variables. The combinations of the quantitative parameter values and qualitative categories suggest many possible treatment conditions.

Insert Table 1 About Here

The variables identified in Table 1 are independent of the behavioral level of a given task or of a given test situation. For each level of behavior (i.e., classification, rule-using, higher order rule-using), it is possible to define an expository or an inquisitory presentation that uses either generalities or instances; it is possible to define generality scope, instance scope, attribute matching, and instance probability; and it is possible to define correct answer, attribute isolation, or algorithms.

Presentation Form

Presentation form deals with the questions: What is presented (generality or instance)? And how is it presented (expository or inquisitory)? Presentation mode indicates whether or not a response is solicited from the student. If the display presents information, it is expository; if it asks a question, it is inquisitory. Presentation content refers to whether the presentation consists of a general statement (generality) or a specific example of a more general class (instance). If the information is an abstraction that refers to the whole class, or relates several classes of specific instances, then it is a generality. In concept instruction, the generality is the definition of the concept. In rule using the generality is the statement of the principle or rule being taught.

Instances are specific examples or members of the class under consideration. In the case of the concept, the instances are specific members grouped by the concept label. In a principle task, the instances are those specific problems that consist of a specific instance from the stimulus class to which the operation has been or can be applied to produce a specific instance of the response class. In problem situations, the instances are those sets of problems for which an operation can be determined by use of a higher order rule. A negative instance is any specific example that is not a member of the class or classes under consideration. For concepts, instructionally relevant negative instances are potentially confusing examples of similar concepts. For principles and problems, negative instances consist of inappropriate applications of the operation or higher order rule.

Any particular instance of a complex cognitive instructional presentation must always feature some combination of presentation mode and content. Combining the qualitative values into a two-way table produces the following presentations: expository-generality (EG), expository-instance (Eeg); inquisitory-generality (IG), and inquisitory-instance (Ieg). In an expository -

generality (EG) presentation, S is presented definitions or rules and directed to study or read these generalities. In an expository-instance (Eeg) presentation, S is presented exemplars or non-exemplars of a given concept along with some indication of appropriate class membership or a sample of the application of a rule to a particular problem. This presentation may be in the form of questions as long as the answer is given at the same time. In an inquisitory-generality (IG) presentation, the student is asked to reproduce or deduce the definition or rule. If this presentation follows an expository-generality presentation, then it requires discriminated recall, not classification, rule using, or higher order rule using. However, if this presentation follows a series of expository-instance presentations, and S is asked to deduce the rule from these previous presentations, then it does require classification behavior above. In an inquisitory-instance (Ieg) presentation, S is presented an exemplar or nonexemplar of a concept and asked to indicate class membership; or he is given a problem and asked to apply the rule.

Simple examples of these combinations are not easy to find, since most strategies are really some aggregate of two or more. In most of the studies cited for illustration, a comparison of presentation form variables was not the primary research question. Since the task variables are independent of the type of complex cognitive outcome, one should be able to identify an instance of each of the strategies as it applies to concept/ classification, principle/rule using, and problem/higher order rule-using tasks.

<u>Concept: Eg-Eeg or Ieg.</u> Tennyson, Woolley, and Merrill (1972) taught the concept "trochaic meter" using a combination procedure consisting of an expository-generality (EG) followed by an expository-instance (Eeg) presentation. The EG presentation consisted of one page containing the following definition:

> Part of the rhythm of a poem is determined by the time between stresses occupied with unstressed syllables or pauses. Denoting the stress patterns is to establish the meter. One of the major meter scansions is named trochee and consists of a stressed syllable followed by an unstressed syllable (marked thusly:)...

The expository-generality presentation was followed by a series of four expository-instance presentations, each consisting of a page containing four short passages of poetry: two exemplars illustrating trochaic meter and two nonexemplars. The instances were clearly labeled as "example" and "not an example." A sample page is as follows:

divergent	matched[1]	Example	Out of childhood into manhood Now had grown my Hiawatha (Longfellow)
		Not an Example	Come to the crag where the beacon is blazing Come with the buckler, the lance, and the bow (Scott)
	matched	Example	Pansies, lilies, kingcups, daisies. (Wadsworth)
		Not an Example	Motherly, Fatherly, Sisterly, Brotherly! (Unknown)

Woolley (1971) replicated this study but modified the instance presentation so that rather than being told which were the exemplars and which were the exemplars and which were nonexemplars, S was required to indicate his choice after which he was given the correct answer.

Concept: EG-IG. As part of a concept study, Johnson and Stratton (1966) included an expository-generality (EG) or "definition" treatment in which each S was shown a page containing four definitions and asked to rewrite the definition in his own words. The inquisitory generality (IG) test was administered nine days later and consisted of a page containing four words which S was asked to define in writing. Because this test does not meet the conditions necessary for classification behavior, it is definition recall.

Concept: Eeg or Ieg-IG. Laboratory studies of concept learning typically present a set of instances--exemplars and nonexemplars--in either an expository or inquisitory mode. When he can, S is asked to indicate the relevant attributes (inquisitory-generality). Clark (1971) reviewed 250 such studies.

Rule: Eg-Eeg-Ieg. In a task designed to teach the rule for winning a NIM game, Scandura, Woodward, and Lee (1967) used three types of presentation: an expository-generality, an expository-instance, and an inquisitory-instance. A NIM game involves two players. The first chooses a number within a limited series (e.g., 1-6); the second player names another number, also within the set; the first player then names another number. This process continues until the cumulative sum reaches or surpasses a given amount. The player who names the final number required to reach the predetermined sum wins the game.

The expository-generality presentation consisted of a rule for winning the game whenever you are allowed to make the first choice. The following page was an expository-instance presentation which consisted of a sample game with the responses of both players indicated. This sample game showed how the rule was applied to a (6, 31) game where the selection set was the digits 1-6 and the final sum was 31. The third presentation was an inquisitory-instance presentation in which S was asked to complete an additional example of a (6, 31) NIM game.

Rule: Eeg-IG. An inquisitory-generality (IG) presentation may be described using the materials previously defined. If the rule for winning a NIM game as previously described by Scandura, Woodward, and Lee (1967) were to be deduced from an expository-instance (Eeg) presentation consisting of a series of differently solved NIM games, then if S can state the rules for winning (IG) his behavior would require rule-using.

Rule: Ieg. Guthrie (1967) used a pure inquisitory-instance (Ieg) procedure involving a rule for solving a particular cryptogram (words formed by scrambling letters). Without telling S the rule, he presented a series of cryptograms, each scrambled according to one rule until S was able to correctly say eight successive words. S was never required to verbalize the rule.

Interdisplay Relationships

The second major class of variables indicated in Table 1 is interdisplay relationships. A display is defined as that material constituting a single presentation which may take any of the four forms previously identified--

expository-generality, expository-instance, inquisitory-generality, or inquisitory-instance.

Interdisplay relationships refers to similarities and differences that relate one display or its elements to another display or its elements. The variables described deal with relation of generalities to generalities, generalities to instances, instances to instances, and instances to the ability of a given S population.

Generality Scope. Generality scope relates generalities to generalities. In complex cognitive behavior, there are at least two types of generalities: concept definitions and rules. A concept definition is basically a list of relevant attributes used to indicate class membership. This list can take several forms: conjunctive, disjunctive, relational, and biconditional (Clark, 1971).

One definition is more restricted than another when the set of objects or events referenced by the restricted definition are a subset of the objects or events referenced by the more general definition. The structure of language is such that almost all concepts are subsets of more general concepts while at the same time, consisting of more restricted concepts. A general concept definition is that of "metered foot."

> Meter involves a measured, patterned arrangement of syllables according to stress and length. A group of syllables which consist of the basic unit of meter is called the foot.

A restricted definition would define a particular kind of metered foot, e.g., trochee.

> Trochaic meter involves a measured, patterned arrangement of syllables where the first syllable is stressed and the second is unstressed. Each set of a stressed followed by an unstressed syllable is a trochee foot.

A rule consists of two sets (concepts), a stimulus set or domain (D-set), and a response set or range (R-set). The rule relates these two sets by means of an operation such that applying the operation to a member of the stimulus set produced one and only one member of the response set (Scandura, 1966, 1968). When defined in this way, rules can be easily ordered as to generality. One rule is more general than another if it includes all of the stimulus elements and their resultant response elements plus additional stimulus response pairs of its own (Scandura, Woodward, & Lee, 1967).

In a previously cited study, Scandura, Woodward, and Lee (1969) investigated extra scope transfer in a rule-using task involving the NIM game. The specific winning rule which applied only to (6, 31) games was as follows:

> "In order to win the game, you should make three your first selection. Then you should make selections so that the sums corresponding to your selection differ by 7."

The more general rule which applies to all (n, m) games was as follows:

> "In order to win the game, the appropriate first selection is determined by adding one to the largest number in the set from which the selection must come and dividing the desired sum by the result. The remainder of this division is precisely the selection that should be made first. Then, you should make selections so that the sums corresponding to your selections differ by one greater than

> the largest number in the set from which the selections must come."

As we observe, the specific rule is restricted to a single type of NIM game, while the general rule refers to all (6, 31) games plus any other NIM games of the form (n, m).

Instance scope. Instance scope relates instances to generalities. When instances possess the attributes listed by the definition, they are said to be a within-scope exemplar. If the instance possesses attributes not listed in the definition, then the instance is extra scope. By definition, nonexemplars are extra-scope instances.

Both of the following passages are examples of metered verse. The foot is underlined.

> Out of Childhood into manhood
>
> The God of love my shepherd is,

In relationship to the definition of metered foot given above, both are within-scope instances. In relationship to the restricted definition of trochaic meter, the first is a within-scope instance, but the second is an extra-scope instance, since it involves Iambic meter (first syllable unstressed with second syllable stressed).

An instance of a rule is a problem that consists of a stimulus situation to which an operation can be applied to produce a member of the response set. Scandura indicated that concepts were special cases of rule-using where the operation was a mapping operation and there was only one member in the response set (Scandura, (1966c, 1968). In the simplest case, an extra-scope problem is one which does not belong to the stimulus set referenced by the rule, but does belong to a more general set that can be encompassed by a more general rule. In the case of higher order rules, where there are a class of operations involved, an extra-scope instance is one that requires a variation of the operation included in the statement of the rule. In other words, the extra-scope instance might be within the scope of a higher order rule, but is extra scope for the more restricted rule that has been stated.

In the previously cited study of Guthrie (1967), Ss were tested using both within - and extra- scope instances. Given two rules for solving cryptograms, one of which was, "Exchange the first and last letters...," each S was tested on within-scope instances which could be unscrambled using one of the two rules he had learned. Second, he was asked to solve "near transfer" items, scrambled according to the same type of rule, i.e., transpositional, but where different letters were transposed, e.g., "Reverse the order of the first half and last half of the letters." Next, he was given "remote transfer" items, requiring application of a different type of-rule, i.e., substitutional, e.g., "Replace the numbers with the correct vowel (a, e, i, o, u equals 1, 2, 3, 4, 5)." Both of the latter "transfer" tasks were extra scope instances.

Attribute matching. Attribute matching relates instances to other instances. Class membership of a given instance of a concept is determined by relevant attributes. All other attributes are said to be irrelevant. Two instances are matched when their irrelevant attributes are as similar as possible and divergent when the irrelevant attributes are as different as possible

(Tennyson and Woolley, 1972). Exemplars may be matched or divergent with other exemplars or with nonexemplars. Obviously, attribute matching is a matter of degree.

Attribute matching was one of the variables manipulated in the previously cited study by Tennyson, Woolley, and Merrill (1972). The four instances on each display consisted of two exemplars of trochaic meter, each matched to a nonexemplar and divergent from each other. The previously cited display from this study illustrates the matched and divergent relationships. The irrelevant attributes included rhyme, number of lines, and subject matter. The first pair are matched in number of nonrhyming lines and basic subject matter. The second pair both consist of four similar words and one line. The difference between the exemplar and the nonexemplar in each pair is the relevant attribute, i.e., type of meter. The two examples are very different (divergent) from each other in subject matter, number of lines, and the fact that one tells a story (or part of one) and the other is merely a set of similar words.

Attribute matching in rule using tasks is much more complex than for classification tasks. Attribute matching can be defined for both the stimulus set and the response set. Traub (1966), without so identifying, introduced a divergent relationship between members of the stimulus set which he called problem heterogeneity. His task was graphical integer addition using a number line. The irrelevant attributes consisted of such things as the magnitude of of the digits involved, the segment of the number line contained in the problem, whether the segment contained a zero point and the direction the arrow was to be drawn.

Thatcher (1972) in an extension of the Traub study, introduced, in addition to the divergent relationship between members of the stimulus set, a matched relationship between exemplars and nonexemplars of the response set. During the presentation, he paired incorrectly solved problems with correctly solved problems. The incorrectly solved problem was matched to the correctly solved problem in the line segment used and the number of numbers involved.

Guthrie's (1967) cryptogram task provides a good illustration of the three levels of attribute matching possible within a problem-solving task. First, instances within the stimulus set (cryptograms to be solved) can be matched or divergent on such dimensions as the number of letters, syllables, and type of word (noun, adverb, and verb). Members of the response set can be matched or divergent on such attributes as number of letters substituted or transformed, or the number of different steps required to unscramble. In both of these sets, these relationships can be between instances, i.e., cryptograms with cryptograms in the stimulus set; correct solutions with other correct solutions in the response set; or between instances and noninstances; i.e., cryptograms with noncryptograms; correct responses with incorrect responses. In addition to matching within and without the stimulus and response sets, one can describe matching relationships between operations. These relationships can also be operations matched or divergent from other operations or operations matched or divergent with incorrect or inappropriate operations.

Instance difficulty. Given a set of objects or events and the directions to identify class membership, students will almost always find some members of the set more obvious than others. Inferences from concept research would indicate that this is a function of the number of relevant attributes defining a given concept (cf. Clark, 1971, statement number 4, twelve studies support this finding), the number of irrelevant attributes present for a given instance of a concept (cf. Clark, 1971, statement number 11, forty studies support this finding), and variances in the difference between relevant and irrelevant attributes (cf. Clark, 1971, three studies support this finding). An operational procedure for assessing the relative difficulties of the instances to be used in a given task was defined by Tennyson and Woolley (1972) and elaborated by Tennyson and Boutwell (1972). The procedure involves an expository presentation of a definition followed by an inquisitory presentation of a large sample of instances, both exemplars and nonexemplars, to a sample of students representative of the target population to be taught. The percentage of the sample correctly identifying a given instance provides an empirical measure of the ease with which a given instance will be recognized. Applying this procedure to the concept trochaic meter, using college Ss yielded distributions which approximate the normal curve (Merrill and Tennyson, 1972a, 1972b; Tennyson, 1972).

In rule using, the variable, instance difficulty, must be defined for both the stimulus set and the response set. Assessing the difficulty of the stimulus set would be similar to the instance probability procedure employed for concept tasks. The difficulty of response set instances most likely depends on similar parameters to that of a concept set; however, in this case some of the irrelevant attributes also consist of the number of logical error possibilities, and the complexity of the response. Perhaps the probability of a given response set instance can also be assessed by an empirical procedure, but such a procedure has not been employed as far as the reviewers have been able to determine.

Higher order rules also involve a set of operations. Just as with an ordinary concept set, this set will also consist of some operations which are more complex, contain more irrelevant attributes, and are more difficult to discriminate, than are other operations in a given set. This introduces the possibility of an operations instance probability. This variable will no doubt prove to be an important variable for the study of problem solving. The reviewers were unable to identify any examples where operations instance probability was applied to problem-solving tasks.

Mathemagenic Information[2]

Mathemagenic information is so called because it is additional or augmenting information provided to facilitate learning. Mathemagenic information is always provided in relation to instances and will only be present when one uses an expository-instance or the inquisitory-instance presentation. There are at least three types of mathemagenic information: correct answer, attribute isolation, and algorithms.

Prompting and feedback are two classes of mathemagenic information variables. Prompting has been called hints, cues, and clues. Feedback has been called reinforcement (probably incorrectly, cf. Anderson, 1967), confirmation, or knowledge of results. For this taxonomy, they are lumped together because they consist of the same information, given at different points in an instructional sequence. Mathemagenic information contained in an expository-instance presentation which precedes an inquisitory-instance presentation is presented before S is asked to respond and is a form of prompting. Mathemagenic information which is presented following an inquisitory-instance presentation, afte the student has responded, is always a form of feedback.

Correct Answer. The simplest form of mathemagenic information consists of providing S with the correct response for a given instance in an expository presentation (prompting) or providing him with the response he should have made following an inquisitory presentation (feedback). In some feedback situations, rather than being told the correct answer, S is merely told he is "right" or "wrong."

Correct answer prompting in concept learning situations is illustrated in the previously cited study by Tennyson, Woolley, and Merrill (1972). After receiving a definition, S was presented examples and nonexamples of the concept trochaic meter. Each verse was accompanied by the word, "example," or "not an example."

Correct answer feedback is illustrated in the same concept task in the previously cited study by Woolley (1971). Woolley did not label the verses of poetry, but asked S to indicate the example. After S had responded, then he was given feedback consisting of the verse labeled "example" or "not an example."

Part of the expository-instance treatment previously described for the Scandura, Woodward, and Lee (1967) study consisted of a correct answer prompt for a rule using task. After an expository presentation of the rule, S was shown a completed example of a NIM game. This is essentially a problem with the correct answers given.

This same study and treatment can be used to illustrate a correct answer feedback procedure for the rule using task. After studying the completed game, S was given a different (6, 31) NIM game and asked to complete the responses. After he had finished, he turned the page and found the game worked for him--or, more precisely, found the correct answers given in a feedback position.

Isolation of attributes. When some procedure focuses attention on the relevant attributes of a given stimulus situation, thus facilitating discrimination, such a procedure isolates attributes. Attributes can be isolated prior to or after a student responds by such devices as underlining, color emphasis, exploded drawings or pictures, italics, arrows, or boxes.

While prompting has been investigated with memory tasks, very little has been done in teaching concepts. Merrill and Tennyson (1972) investigated a prompted expository presentation of the trochaic meter task which employed an attribute isolation technique. They defined trochaic meter and the use of stress symbols and then marked the rhythm pattern in each verse

(/ for stressed, ◡ for unstressed). A prompted verse appeared as follows:

/ ◡ / ◡ / ◡ / ◡
Out of childhood into manhood

/ ◡ / ◡ / ◡ / ◡
Now has grown my Hiawatha.

Without identifying it as such, Roughead and Scandura (1968) used an attribute isolation prompting technique for teaching a rule using task. The task was to "discover" the general formula for summing number series. The presentation consisted of three series and their respective summing formulas. The problems were presented in tabular form which isolated the term numbers, term values, and cumulative sums. All of the attributes are present or can be determined from the unprompted presentation, but are more obvious in isolated form. If, in this task, S had been given the series and asked to find the formula first, then presenting the tabular presentation would constitute feedback.

Algorithms.[3] An algorithm is a rule for analyzing a given instance in order to identify its class membership or to set it up for applying the operation (search strategy). It may also be a rule for synthesizing an answer or product by applying an operation (production strategy). A search strategy in a concept task is a rule for searching for and identifying the relevant attributes. In a rule using task, this search strategy is extended to rules for identifying whether or not the term is one to which a given operation applies or rules for applying the operation to the stimulus term. Production strategies are rules for producing the response term or a product. Since algorithms are themselves generalities (rules), it is accurate to say that they are rules for solving a given problem. In the trochaic meter task previously described, an algorithm (search strategy) for identifying relevant attributes might be as follows:

1. Write or imagine the line broken into syllables as follows:

   "out of child hood in to man hood . . ."

2. Read the line aloud listening for stressed syllables.

3. In trochaic meter, the odd numbered syllables are stressed and the even unstressed, etc., for other patterns of meter.

Scandura and some of his students (1966) employed an algorithm prompting technique in teaching a rule-using task. The stimulus set was 4-tuple sets of integers, e.g., (4, 8, 9, 3); the response set consisted of an integer which could be derived uniquely from three numbers in the set by using two arithmetic operations, e.g., add the first and second integers in the set and subtract the fourth. The task for the student was to "discover" the operation which applied to a given set of items. Given this information, a general algorithm can be formulated:

> After indicating that there is a rule . . .
>
> To help you discover this procedure as rapidly as possible, you should try to determine the three specific positions in the 4 - tuple and an arithmetic rule involving two operations (e.g., add and divide) which combines the numbers in these positions to yield the corresponding number.

In a sense, providing an algorithm for any task makes it a rule-learning task by providing instruction within instruction. All previous variables applying to rule learning can also be applied to providing an algorithm as a prompt or as feedback.

Merrill (1965, 1970) provided algorithms in a feedback position in an imaginary scient task. This treatment was labeled "specific review" and consisted of a step-by-step solution of the problem which the student had previously tried to solve.

Variable Parameters - Sequence, Quantity, Pace. Each of the quantitative variables listed in Table 1 can be applied to any of the qualitative variables previously described. The next few paragraphs will define a few possible combinations to clarify the meaning of the quantitative parameters.

Sequence. Sequence deals with what should be presented first (order), how reoccuring procedures should be patterned (schedule), and whether different kinds of events should be presented as successive or simultaneous displays.

Sequence Order. Prompting and feedback are one combination of order and the mathemagenic information qualitative variables already discussed. Most of the treatments described to illustrate mode and content of presentation were also combination treatments ordered in a variety of ways: Eg-Eeg or Ieg (Tennyson, Woolley, and Merrill, 1972; Woolley, 1971); EG-IG (Johnson and Stratton, 1966); Eeg-Ig or Ieg-IG for an inductive concept presentation (Clark, 1971); and EG-Eeg-Ieg in a rule instruction task (Scandura, Woodward, and Lee, 1967). The primary variable investigated in each of these studies, however, was no order crossed with presentation form.

Sequence Schedule. Schedule refers to the pattern of receiving qualitative variables. One example deals with whether an expository generality (EG) followed by inquisitory-instance (Ieg) presentation which includes several concepts or rules, should be presented Eg Eg Eg-Ieg Ieg Ieg or whether an Eg Ieg-EG Ieg - EG Ieg schedule would be better. Another schedule question deals with interdisplay relationships. For example, Tennyson, Woolley and Merrill (1972) described three instance probability schedules in an expository-instance presentation of instances of trochaic meter. Each treatment involved four displays, each containing four instances. One probability sequence was high-medium-medium-low; a second was high-high-high-high; a third sequence was low-low-low-low. This study did not provide a direct comparison of these variables because each was confounded with attribute matching which also varied across treatments. Feedback and prompting schedules are also applications of this quantitative variable.

Sequence Simultaneous versus sequential presentation. This variable is in one sense a special case of scheduling, but sufficiently important and different that it has been separated for discussion. Questions which result from applying this quantitative dimension to the qualitative variables already described, are as follows: Is a display consisting of both an expository-generality and an inquisitory-instance presentation, better than presenting the inquisitory-instance presentation following the expository-generality presentation? Similar questions could be asked about the other possible combinations or about interdisplay relationships, e.g., in attribute matching, should matched instances be presented simultaneously or sequentially? Other simultaneous versus sequential questions could deal with various combinations of mathemagenic information, and various combinations of mathemagenic information with interdisplay relationships.

Quantity. The question of how many or how much can be applied to almost every value of each of the qualitative variables identified. It is likely that many such questions are specific to individual students; nevertheless, there are limits which probably apply to most students.

A second question related to quantity is the relative proportion of the components of various combinations. What should be the ratio of instances to generalities; the ratio of exemplars to nonexemplars; the ratio of high to low probability instances; or the ratio of prompting to feedback?

Pace. In the realm of complex cognitive behavior, very little has been done to investigate this parameter. The basic question deals with programmed paced versus self-paced presentations. This is usually applied to the overall presentation, but it is possible to introduce paced segments into a primarily learner paced presentation. Other decisions need to be made about whether mathemagenic prompt information should be limited in its availablility? Or feedback? Do some paced presentations facilitate learning?

As soon as pacing becomes a factor, then interval time becomes critical. The intervals may be very short, such as a few seconds between successive displays, or long, such as several days or weeks. Delay is a critical part of the definition of retention and, hence, whenever retention is examined as a dependent variable, one must be concerned with the retention interval.

Experimental Research on Task Variables Promoting Higher Cognitive Behavior.

The following paragraphs review some of the best research on task variables done to date using concept, principle and problem tasks. While it would be presumptuous to say this review is exhaustive, it is nevertheless true that there are very few continuing systematic research efforts in this area. Several criteria were used in selecting studies for inclusion. First, they must involve at least classification behavior of above. If a task involved concept, principle or problem content but the behavior required was discriminated recall, then the study was not included. For example, Kulhavy and Anderson (1972) reviewed a series of studies on delay of retention with multiple choice tests. In every case, the dependent variable was performance on a second administration of the test after the correct answers had been given following the first administration. Even though the tasks were classed as "meaningful verbal materials," sometimes involving concepts and principles, the behavior required was instance recall. Most of the investigations of prose learning (See Glaser and Resnick, 1972 for review) are not included, even though the content sometimes involves concepts and principles because the behavior required is usually recall.

A second criterion was fidelity to classroom or "real world" content. Clark (1971) reviewed more than 250 concept studies. While the findings of these studies have relevance for task variables, most were highly structured laboratory tasks. As indicated in Clark's (1971) paper, real world concepts often consist of an indefinitely large number of possible instances; laboratory concepts are usually limited to a small number of instances. Real world instances are frequently abstract and become more so as a student goes from grade to grade; laboratory concepts are usually concrete physical objects. The attributes of real world concepts are themselves properties (concepts) and, hence, each such property can be viewed along several continuous dimensions; the attributes of most laboratory concepts are dimensions which have two or more discrete values. The number of attributes, especially irrelevant attributes, and their possible combinations is indefinitely large for real world concepts; the number of dimensions for laboratory concepts is small as are their possible combinations.

Research on task hierarchies

Previous research on task hierarchies (See Briggs, 1968; Gagné and Rohwer, 1969; Glaser and Resnick, 1972, for reviews of this research) has failed to distinguish horizontal from vertical hierarchies. A horizontal hierarchy is one within a single level of learning, i.e., a hierarchy consisting entirely of concept-classification or principle-rule-using tasks. The Gagné and Merrill taxonomy (Gagné, 1970; Merrill 1971a/1971b) suggests a vertical hierarchy in that there is always some discriminated recall behavior prerequisite to any classification behavior, some classification behavior prerequisite to some rule using behavior, etc. With one exception (Wang, Resnick, and Boozer, 1971), all of the studies cited by Glaser and Resnick (1972) are horizontal hierarchies. Wang, Resnick and Boozer (1971) investigated skills in a math curriculum and found that if S can match an integer to a set of an appropriate size (concept-classification), that he can also read the integer name shown the figure (paired associate-discriminated recall). However, this study did not demonstrate that S can name numbers without being able to match this symbol with a given set.

Sudweeks and Merrill (1971a) investigated misconception errors on a defined concept resulting from misconceptions present in the attribute concepts. The task was to identify unencountered nonsense words which are "frams." A fram is a single syllable word ending in a simple consonant, preceded by a single vowel, preceded by one or more consonants. Ss were pretested on their ability to identify consonants and vowels. All Ss who already knew when w and y were vowels, that qu represents the consonant blend kw, or that x represents the consonant blend ks, were eliminated from the study. Ss in different groups were taught none, one, or a combination of the above attribute exceptions and then all Ss were taught the concept "fram" using an expository-generality followed by an inquisitory-instance presentation. All Ss who were taught none or only one exception made misconception errors completely consistent with the attribute misconceptions which had not been corrected. Correcting one of the possible misconceptions did not correct for other misconceptions possessed by a given S. This study provides some evidence for a vertical hierarchy.

Scandura (1966a) investigated the effect of different kinds of prerequisite training on the ability to use an algorithm (rule) to solve a principle rule-using task. The D-set consisted of problems made up of cards containing two geometric figures (circles, triangles) which can each have two values (e.g., black or white) on three dimensions (color, size, shape), where each card is labeled +all, -all, +1, etc., to indicate the number of relevant symbols in an associated symbol set (see Scandura, 1964). The R-set consisted of symbol sets where one or more can be associated with all three cards in a given problem. The operation was left unspecified but all Ss were taught to use an algorithm (search strategy) for determining the associated symbol set (s). Treatments varied the kind of prerequisite training. This prerequisite material consisted of three tasks: a paired associate task, where the behavior was discriminated recall of the symbols (S_1, L_1, C_2, etc.) associated with the attributes of figures on the cards; a concept task where the behavior was to classify instances of "+all" symbol sets; and a second concept requiring the classification of "+1" symbol sets. The symbol group (S) was taught the discriminated recall of the symbols; a second group (SC), in addition to the material taught to group S, was taught the "+all" and "+1" concept tasks. A third group (P) was taught all the material shown to groups S and C and was also shown sample problems. A control group was given no pretraining. All groups were then taught an algorithm (search strategy) for one type of problem. The dependent variables were the attempted solutions of six problems; two within the scope of the algorithm, two within the scope, but requiring slight variation in the algorithm, and two extra scope problems requiring considerable variation in the algorithm. On the four within scope problems, groups S, SC, and P identified significantly more correct sets than did the control group. On the two extra scope problems, group S did not do any better than the control group; however, groups SC and P correctly identified more symbol sets than group S. There was a significant drop in performance for all the experimental groups on the four problems involving either slight or considerable variation in the algorithm.

In a more extensive second study, Scandura (1966b) again varied the amount and kind of prerequisite material. In addition, he also varied sequence and amount of practice. The dependent variables included performance (number of correct sets and time) on seven practice problems as well as performance on two extra scope transfer problems. Groups which were pretrained on the paired associate symbol task and the "+all" concept task did significantly better on the rule-using task and the extra scope problems than did groups which were not pretrained. Sequence and practice did not have marked effects compared with the effect produced by the kind of pretraining. Both of these studies lend considerable support to the necessity of acquiring relevant paired associate-discriminated recall and concept/classification behavior before undertaking principle/rule-using tasks.

Much more needs to be done on the study of vertical hierarchies. Scandura (1966 a, b) did not systematically investigate all of the prerequisite concepts in his task. He did not demonstrate that Ss could have the concept and still be unable to use the rules. Sudweeks and Merrill (1972) capitalized on widespread, overlearned misconceptions. Would the same results follow in teaching a new concept-principle hierarchy where the concepts were new to the student and where the rule was taught but not all of the component concepts? If the hypothesis of the content-behavior hierarchy is to be established, it must be demonstrated that Ss can demonstrate adequate classification of component concepts but be unable to use the rule; be able to use some rules, but still not determine the operation for new D-sets. Questions such as the following have not yet received adequate investigation: Do the columns in Figure 1 represent prerequisite behaviors for each content-behavior combination on the diagonal in a given column? What are the prerequisite discriminated recall behaviors for concept-classification-recall of the definition (relevant attributes) or recall of encountered instances? Can it be demonstrated that S can possess the prerequisite but not the subsequent behavior in a given column? Are the behavior-content categories exhaustive or do other categories or subdivisions need to be introduced?

Presentation Form

Generalities Alone versus Generalities Plus Instances. Several studies (Guthrie and Baldwin, 1970; Merrill and Tennyson, 1972b; Anderson and Kulhavy, 1972; Watts and Anderson, 1972) have compared combinations of EG, Eeg, and Ieg presentations. The general conclusion is that being presented examples (Eeg) or asked to identify examples (Ieg) following a presentation of the definition is superior to a presentation of either the definition or examples alone.

Guthrie and Baldwin (1970) taught intercity fifth graders the concepts "a" before a consonant word and "an" before a vowel word. Groups differed in that one group was shown and asked to recall the definition (EG-IG). The other group was also given practice in identifying instances (EG-IG-Ieg). On tests requiring S to recognize and produce the correct form, there was no difference for high ability Ss, but low ability Ss did better after practice at classification.

In a concept task teaching trochaic meter to college Ss, Merrill and Tennyson (1972b) compared various types and combinations of presentations. Treatments included definition only (EG); examples and nonexamples (Eeg); and definition followed by examples and nonexamples (EG-Eeg). The combination treatment Eg-Eeg was superior on correct classification to either presentation in isolation. EG or Eeg in isolation are not significantly different from each other, but both produced significantly more over-generalization errors than did the combination.

Anderson and Kulhavy (1972) presented Ss definitions of infrequently appearing words. One group was required to use the word in a sentence (Ieg), while the other group read the definition three times (EG). The test consisted of multiple-choice questions where the stem was the concept name and the alternatives were descriptions of specific positive and negative instances. Ss who were required to use the word in a sentence performed better on the classification test than did Ss who merely read the definition aloud three times.

Watts and Anderson (1971) conducted a prose learning study which compared three types of inserted questions: questions requiring S to identify an unencountered example of the concept when given its name (classification), questions requiring identification of an example previously used in the text (instance recall) and questions asking S to identify the scientist associated with a given concept (discriminated recall). There were five passages each defining, presenting two illustrations, and identifying the scientist associated with the concept. One of the three types of questions followed each passage. The posttest consisted of all three types of questions. The classification group did better on all types of questions than did the other groups. The group with inserted discriminated recall questions did worst. The classification group also took less time to complete the test than the other groups.

The studies cited do not adequately compare the main effect variables of expository versus inquisitory or generality versus instance. Most studies in the literature are comparing combination treatments which make the adequate determination of the relevant contribution of these variables difficult. The limited number of studies where inferences can be drawn needs to be replicated for other tasks and with other populations before we would be willing to state propositions including these variables. Often these studies (e.g. Guthrie and Baldwin, 1970 and Merrill and Tennyson, 1972b) are confounded in that one treatment includes more material than another and consequently allows for more practice. Additional research is needed of the following type: (1) Direct comparisons of expository versus inquisitory and generality versus instance in two-way designs which include adequate controls for increased practice, number of examples, and use of nonexamples. (2) Sequence studies in which one variable is presence or absence of the generality, a second variable is expository-inquisitory sequence (Eeg Ieg versus Ieg Eeg) and a third variable is repetition of the same instance versus practice with different instances. (3) Multiple generality studies which compare presenting the generalities first and then instances for the whole set (EG_1 EG_2 EG_3 eg_1 eg_2 eg_3) versus presenting instances following each generality (EG_1 eg_1 EG_2 eg_2 EG_3 eg_3). Studies would need to compare both expository and inquisitory modes for presenting examples and repetition of a single example versus several different examples. (4) Studies comparing the minimal instance set necessary to logically eliminate irrelevant attributes with a student selected set. These studies should look at aptitude variables which may interact with the instance set chosen. (5) Studies investigating the ratio of examples to nonexamples. (6) Studies which allow the learner to indicate what he wants to see next, a worked example (Eeg), a practice example (Ieg), or a generality (EG) as well as when he wants to see or try another example compared with linked controls. (7) Studies which compare learner-paced inspection time with group or machine paced times. There are studies in the literature which investigate some of these variables in memory level tasks but almost no research on higher cognitive behaviors.

<u>Interdisplay Relationships</u>

<u>Generality and Instance Scope</u>. A series of experiments (Craig, 1956; Kersh, 1958; Haslerud and Meyers, 1958; Gagné and Brown, 1961; Kersh, 1962; Wittrock, 1963; Guthrie, 1967; Worthen, 1968) investigated "discovery learning"

using principle/rule using tasks consisting, in most studies, of either rules for summing number series or rules for deciphering cryptograms. The discovery treatment was usually some form of inquisitory instance (Ieg) presentation with or without mathemagenic information in the form of hints or attributes isolation. This IEG treatment was compared with an expository-generality (EG) presentation of the specific rules required for solving one class of problem. In most studies, Ss were tested with similar (within-scope) problems and also with problems of the same type, both which required a different rule than the one taught (extra scope problems). The usual finding is that, when compared to Ss receiving an Eg-Ieg presentation, Ss receiving some form of Ieg presentation first do worse on within scope problems, especially on an immediate test; but do better on extra scope problems, especially after a delay interval.

In a problem-solving task where different number series comprised the set of D-sets and where the operations set is comprised of the summing formulas for each series, Roughead and Scandura (1968) compared the expository presentation of a general derivation rule for determining the appropriate formula for a given series (EG_1) with the expository presentation of specific rules indicating the summing formula for a given series (EG_2) and with an inquisitory procedure requiring the student to derive the formulas for several series (IG_2). The dependent measures were weighted scores combining the time and the number of hints required to derive specific rules for unencountered within-scope and extra-scope problems. On both the within-and extra-scope problems,

EG_1 and IG_2 were not significantly different from each other, but both groups performed more efficiently than EG_2 or a combination treatment where IG_2 followed EG_2. It was concluded that what is learned in the discovery (IG_2) situation is the general derivation rule (G_1) and that an expository presentation of the specific rules (EG_2) prior to discovery (IG_2) limits the student's tendency to discover the more general rule.

Paul Merrill (1972) taught Ss to use rules of an imaginary science under four conditions: Eeg-Ieg, where Ss were given an example which demonstrated the application of the rule (expository-instance) and were then asked to solve three additional examples of the same type (inquisitory-instance). If they solved two out of three, they were given an example of the next rule. If they did not, they were given another solved instance of the first rule plus three additional problems. The pattern was repeated up to five times for a single rule for ten rules. (EG-rule Eeg) - Ieg Ss were shown a stated rule simultaneously with the first solved example; (EG-Objective-Eeg) - Ieg Ss were shown an objective simultaneously with the first solved example; (EG-both Eeg) - Ieg Ss were shown both the objective and the rule. The dependent variable was the number of examples required to learn the task. Those Ss who received rules required fewer examples and learned the task in less time than those not receiving rules. There was no difference for objectives. Unlike previous research on discovery, however, the rule groups did better on a transfer test. This finding is consistent with Roughead and Scandura (1968) in that when the rule is general and Ss are given practice with a variety or problems then transfer is better for EG group than for Ieg groups.

Scandura, Woodward and Lee (1967) investigated rule and instance scope in a principle-rule-using task consisting of NIM games (a mathematical game where two players alternatively pick numbers from a given set, the first play to pick the number which exactly reaches a predetermined sum wins). The D-set consists of specifications for the number set and winning sum, the operation is a rule which allows the first player to win the game. The investigators compared a general rule (G), which applied to any game, with a partially specific rule (SG), which applied to (6, m) games, with a specific rule (S), which applied only to (6, 31) games. Two control groups were included; one received only a worked example (Eeg), the other received no instruction, only the test. The dependent variable was the number of Ss who used an appropriate procedure on three NIM games where the first was a within-scope instance for all three rules, a (6, 31) game; the second was within-scope for G and SG, but extra-scope for S; the third was within-scope for G, but extra-scope for S and SG. The results show that the S rule was easier to apply than SG or G; that there was no extra-scope transfer; that if S applied the rule he was taught on problem 1, he used the same rule on problems 2 and 3 even though it was inappropriate on the extra-scope transfer problems.

Scandura and Durnin (1968) extended this study by comparing a restricted general rule (G') which is stated in terms of a specific case, with a general rule restricted in one dimension and specific in the other (SG') and with a specific rule (S') which applies to only one specific NIM game. Specific rule Ss were unable to solve any extra-scope problems, two out of 22 SG' Ss solved the extra-scope problems, only two G' Ss were able to solve one of the problems which were extra-scope for the other groups, but supposedly within-scope for them. When a restricted general rule is used, there is some transfer to the general case for some Ss, but not as much as one would hope. Additional research needs to address the question, "What procedures could be used to promote wide spread application of a general rule?"

Roughead and Scandura (1968) and Paul Merrill (1972) pointed out that previous "discovery" studies used a restricted rule and measured transfer on extra-scope problems while they used a general rule and hence were looking at performance on within-scope problems. If these studies had been conceptualized with these variables in mind, much confusion may have been eliminated. Further replication and extension is necessary to determine if an expository presentation of a general rule is always superior to inquisitory presentation of instances. Several questions have not been resolved. (1) Can set function descriptions be applied in the humanities as well as to math science type subject matters? If so, is the expository presentation of a general generality the superior procedure? (2) If a general generality cannot be determined are there alternative ways to facilitate extra-scope transfer? The authors would hypothesize, generalizing from their concept research, that presenting a variety of expository generalities each followed by inquisitory instance practice would result in better extra-scope transfer than the inquisitory instance ("discovery") practice alone.

Scandura and his associates have consistently found that restricted rules are easier to learn but promote less transfer. A challenging research question is how to teach general generalities. Several questions need investigation. (1) Will a restircted to general generalities presentation prove better than presenting the general generality at the beginning assuming one controls for amount of inquisitory instance practice. (2) Is presenting (EG) a variety of restricted generalities each followed by inquisitory instance practice more effective than the repeated presentation (EG) of a more general generality followed by the same inquisitory instance practice? Generality and instance scope has received such minimal research attention that considerable work is necessary before one will be able to state prescriptive propositions that will assist instructional development.

Attribute Matching and Instance Difficulty. Tennyson (1972) taught seventh grade Ss the concept adverb. His treatments manipulated attribute matching and instance difficulty (empirically determined by instance probability analysis). His dependent variables were correct classification (number of unencountered instances correctly identified) and three predicted error patterns: overgeneralization (incorrectly identifying low probability non-examples as examples), undergeneralization (failure to identify low probability examples as examples) and misconception (focusing on an irrelevant attribute thus failing to correctly identify examples not having this attribute and identifying as examples some nonexamples which do share this attribute). The combination treatment producing the best correct classification consisted of examples matched to nonexamples, examples divergent with other examples, and a range of high to low probability instances. Overgeneralization resulted when examples and nonexamples were unmatched and all low probability instances were used. Undergeneralization resulted when only high probability instances were used. Misconception resulted when examples were not matched to nonexamples and examples were convergent on the irrelevant attribute of an -ly ending. All predicted outcomes were significant beyond $p < .001$. In a second experiment, a parallel set of treatments eliminated all nonexamples. In this study, performance on the test demonstrated random responding which did not differ from a control group, which received no instruction. Contrary to findings using laboratory concepts, when appropriately matched to examples, nonexamples are apparently a critical part of instruction on "real world" concepts. In an earlier study, Tennyson, Woolley and Merrill (1972) found the same results using college Ss and the concept trochaic meter. Merrill and Tennyson (1972a) also replicated these results using college Ss and the concept RX_2 crystal structures.

Using a rule using task which taught sixth grade children graphical integer addition, Traub (1966) investigated problem hetrogeneity (divergent instances). The three treatments included a heterogeneous (divergent) problem group, a homogeneous (convergent) problem group, and a control group which worked irrelevant problems. The dependent variable was performance on 26 unencountered divergent problems using both arrow length and correct integer answer as measures. The heterogeneous problem group demonstrated superior performance.

In a problem solving task where the set of D-sets consist of n-tuples (sets of 3, 4, or 5 integers, e.g., 3, 2, and 7); the operations set consists of two arithmetic operations, (+, -, ÷, and x) for combining three of the n integers in an n-tuple; and the response set consists of the resultant integers; Scandura and Voorhies (1971) investigated the effect of the number of irrelevant operations and the number of irrelevant integers (attributes) on instance difficulty. Students were told that two operations were to be used out of two, three, or four possible operations and that three out of three, four, or five integers in an n-tuple were to be combined by these operations. A linear relationship was found in that one irrelevant operation or one irrelevant integer was more difficult than none; both one irrelevant operation and one irrelevant integer were more difficult than one or the other; etc. The effect was more pronounced for irrelevant integers than for irrelevant operations, probably because Ss could predict likely operations from the size of the response terms. Just as the number of irrelevant attributes increases the difficulty of correctly identifying an instance of a concept (see Clark, 1971), so also increasing the number of attributes and/or operations increases the difficulty of correctly solving a problem. This finding represents a way to identify instance difficulty but is a variable that cannot usually be manipulated in teaching real world concept or problem solving tasks.

Markle and Tiemann (1972) investigated various types of definitions and their effect on instance difficulty. Using the concept morpheme they found that a definition consisting of the critical relevant attributes or one which also included technical irrelevant attributes was superior to definitions including nontechnical irrelevant attributes (typical of textbook definitions). Students in a control group who received Webster's definition performed least adequately.

The research cited has exciting implications for instructional design and does suggest the following proposition for concept instruction. <u>If during practice, using either expository or inquisitory instance presentations, examples are divergent from examples and matched to simultaneously or sequentially presented nonexamples then correct classification of subsequent unencountered instances is more probable</u>. Correlaries can be stated indicating specific classification errors resulting from other relationships. While this research is promising, much replication with other concepts and populations is necessary. Some yet unanswered questions deal with interactions of presentation mode and attribute matching. Under what conditions is simultaneous superior to sequential presentation? How does one most effectively teach sets of concepts versus a single concept? Are the examples of one adequate nonexamples of the other?

Attribute matching and instance difficulty have been almost completely neglected in research on rule using and problem solving. Critical questions yet to be resolved include the following: What is a nonexample of an operation in a rule-using task? Is it an incorrect solution? Which should be matched R-set (range), D-set (domain), or operation? What determines difficulty, the

D-set, R-set or operation? For rule using what are equivalent errors to misconception, overgeneralization, and undergeneralization? Considerably more work is necessary before we would wish to attempt propositions in this area.

Mathemagenic Information

Attribute Isolation-Prompting. Using the concept trochaic meter, Merrill and Tennyson (1972b) investigated attribute isolation in a prompted sequence. The procedure was a definition (EG) presentation followed by divergent sets of examples matched to nonexamples (Eeg). In the prompted attribute isolation group, poetry passages were marked with scansion symbols indicating stressed and unstressed syllables and accompanying comments which pointed out discrepencies such as masculine endings. Groups receiving the prompted passages performed significantly better than the unprompted groups on a correct classification test using new unencountered examples. Unprompted presentation of instances tended to produce more overgeneralization errors than the prompted presentation.

Attribute Isolation-Feedback. Young, Smith and Merrill (1972) investigated specific review (attribute isolation feedback) using the trochaic meter concept task. Ss were given the definition (EG) followed by unidentified positive and negative instances (Ieg). After responding, experimental Ss were told whether their answer was correct and were shown the passage with the stress pattern indicated together with an explanation calling attention to unusual patterns. Control Ss were merely given correct answer feedback.

The specific review groups had higher correct classification scores than the controls. The no review groups made more overgeneralization errors.

Algorithms-Feedback. In a series of studies (Merrill, 1965; Merrill and Stolurow, 1966; Merrill, Barton, and Wood, 1970; Merrill, 1971), using a complex imaginary science, the investigators explored algorithm feedback on rule using questions. The task involved a hierarchically related set of principles for predicting speed and position of satellites in a Xenograde system. Problems required rule using in unencountered problem situations. The task consisted of five programmed lessons, each followed by a quiz. In the Merrill, Barton, and Wood (1970) study, specific review was given following each missed question during the lessons. No feedback was given on lesson quizzes. Specific review consisted of a step-by-step presentation of the rule(s) required to solve the problem presented. This represented a restricted presentation of the general rule presented during the instruction. The feedback was complex, also involving attribute isolation on some frames. Specific review following incorrect responses increased efficiency (time spent on each frame in subsequent lesson), while the number of errors made on quizzes and the terminal test remained the same. The overall time required to learn the task was less for specific review Ss even though they were presented approximately 30 percent more material.

Prompting and feedback (K of R) have received considerable attention primarily using tasks which require discriminated recall behavior. However, very little has been done to systematically investigate these variables for tasks which require higher cognitive behaviors. While previous work has concentrated on response prompting or feedback the research cited suggests that mathemagenic information which isolates stimulus attributes is more promising. Many questions remain to be answered including the following: Are prompting and feedback equivalent procedures? Is stimulus directed mathemagenic information more effective than response directed information? Is there an interaction?

Algorithms have tremendous promise as a powerful instructional tool. Work on mnemonics (see Bower 1970; Rowher 1970) in memory tasks are sure to have parallels in higher order tasks. Directions for processing information (search and production strategies) may prove to be the most powerful instructional tools available. More adequate definition is required for such strategies as well as questions dealing with removal of mathemagenic information, prompt-feedback ratios, student requests versus program provision of mathemagenic information, schedules of feedback-wrong, right or both, and other similar questions.

Implications of a Task Variable Taxonomy for Instructional Development

Instructional developers are often unable to profit from the efforts of their colleagues because it is difficult for other instructional developers to unambiguously describe their products or for instructional researchers to adequately describe their experimental treatments. The proposed task variable taxonomy enables an instructional product designed to promote higher cognitive behaviors to be unambiguously described thereby facilitating communication and comparison of instructional treatments. Each segment can be identified as a combination of expository-inquisitory mode and generality-instance content; each generality and instance can be described in terms of their relationship to other generalities and instances within the treatment; and any mathemagenic information can be identified. With the appropriate selection of symbols it would be possible to completely describe a given instructional strategy in a very limited space with less ambiguity than is usually present in pages of description. For example, the following shorthand:[4] Concept: EG, XEG, $3d\uparrow\ E\hat{eg}$, $Ieg|\overline{eg}m$ $\widehat{(-Eeg/\overline{eg}m)}\triangle d\uparrow$. describes a concept treatment which consists, first, of an expository presentation of the definition (EG). This display is followed by further expository explanation of the definition such as paraphrase or analogy (expanded generality XEG). The third display consists of a series of three divergent (d) exemplars (eg) arranged in ascending difficulty order ($\uparrow$) using a prompted ($\wedge$) expository (E) presentation. The fourth display will provide an opportunity for S to identify (inquisitory I) which of a matched (m) exemplar (eg) nonexemplar ($\overline{eg}$) pair is a member of the concept class ($Ieg/\overline{eg}m$). If he makes the incorrect choice (-...) he will be presented a prompted expository comparison of the matched exemplar non-exemplar pair as feedback ($\widehat{-Eeg/\overline{eg}m}$). The delta ($\Delta$) indicates that S has the option to try as many choices as he feels necessary. Each successive matched exemplar and nonexemplar pair will be matched to each other but divergent (d) from the previous pair and in ascending order of difficulty ($\uparrow$). Unless otherwise indicated, the definition is not restricted and all instances are within scope. The above explanation is still shorthand based on the previous descriptions in this paper. A standard, nonshorthand description of this instructional treatment would take several pages and if typical of such descriptions already in the literature would still be incomplete and ambiguous.

A taxonomy of task variables suggests a lesson component approach to instructional development by making the distinction between content components and instructional components less ambiguous. The content for a task designed to promote higher cognitive behavior consists of four files (sets) of information. A generality file consists of a precise statement of each definition, rule, or higher order rule to be taught. An expanded generality file consists of further explanations of each of these generalities and can take the form of restricted generalities, paraphrased definition or rule statements, analogies, and warnings to avoid certain pervasive irrelevant attributes or procedures. Each generality should be accompanied by an instance file that consists of exemplars and nonexemplars of a concept or problems to which the operation of a rule or higher order rule can be applied. To be complete this file should consist of matched exemplar nonexemplar pairs, divergent instances, and a range of difficulty. Each instance file should be accompanied by a mathemagenic information file that consists of each instance or instance pair with the appropriate attribute isolation or algorithm applied. There may be one or more such displays for each instance or instance pair. These four files represent the basic content components of any lesson designed to teach higher cognitive behavior and can be combined into numerous instructional or evaluation strategies.

An evaluation or instructional strategy is constructed by combining the various combinations of the qualitative task variables with the quantitative parameters identified in the task variable taxonomy. A very large variety of such strategies are possible and the taxonomy makes an unambiguous description possible as well as suggesting consideration of variables which otherwise might be left to chance.

Once a strategy has been selected it is necessary to prepare two types of management displays. The first are control displays which provide directions to the student as to how to proceed, the options which are available to him, cumulative information about his performance, and other directions which enable him to interact with the instructional system. The second type of management display is wrap-around that consists of those questions and directions necessary to present a given generality or instance to the student. These wrap-around displays often take the form of patterns in which a number of different instances can be inserted and which provide directions as to how the student is to proceed.

Summary

Five premises necessary for the formulation of instructional theory are stated. As a suggested elaboration of the "behavior classification" premise a two-way (content and behavior) classification scheme for higher order tasks is described. The content dimension includes paired associate, concept, principle, and problem tasks. The behavior dimension includes discriminated recall, classification, rule using, and higher rule using. Each two-way category is illustrated by treatments from recent research on higher order tasks. As a suggested elaboration of the "manipulation of task variables" premise a taxonomy of task variables is described. The taxonomy is divided into qualitative variables and quantitative parameters. The qualitative variables include the major categories of presentation form, interdisplay relationships, and mathemagenic information. Specific variables are identified in each category. Parameters include major categories of sequence, quantity and pace - specific parameters are identified for each category. The variables in the taxonomy are illustrated by treatments from recent research on higher order tasks. Research investigating some of the variables in the taxonomy is described and critiqued. Some additional needed research is suggested. A final section indicates implications of this research and the task taxonomy for instructional development.

Table 1

Classification of Task Variables for

Promoting Higher Cognitive Instructional Outcomes

I. <u>Qualitative Variables</u>

A. Presentation Form

1. Presentation Mode

(E) Expository (telling) - (I) Inquisitory (asking)

2. Presentation Form
(G) Generality (definitions or rules) - (eg) Instance
(positive or negative examples or problems)

NOTE: Complex cognitive instruction is composed of displays, each representing one of the following combinations:

1. Expository Generality (EG)
2. Expository Instance (Eeg).
3. Inquisitory Generality (IG)
4. Inquisitory Instance (Ieg)

B. Interdisplay Relationships

1. Generality Scope (Restricted General)
Relationship of generalities to generalities
2. Instance Scope (Within Scope - Extra Scope)
Relationship of instances to a generality
3. Attribute Matching (Matched - Divergent)
Relationship of instances to instances
4. Instance Difficulty (Easy - Hard)
Relationship of instances to a given population of students

C. Mathemagenic Information

1. Correct answer (Yes - No)
2. Attribute isolation (Yes - No)
3. Algorithm (Yes - No)

NOTE: Prompting is mathemagenic information presented prior to Ss responding to an inquisitory display. Feedback is mathemagenic information presented following Ss response to an inquisitory display.

II. Quantitative Parameters

Any quantitative parameter can be applied to any qualitative variable.

A. Sequence

1. Order (which comes first?)
2. Schedule (sequence patterns)
3. Simultaneous versus sequential

B. Quantity

1. Number (how many?)
2. Ratio (in what proportions?)

C. Pace

1. Learner versus program
2. Intervals

Figure 1. Two Dimensional Task Classification--Behavior & Content

Behavior Dimension		Paired Associate	Concept	Principle	Problem
	Higher Order Rule Using				Problem Solving Higher order rule using
	Rule Using			Analysis (rule learning) Rule Using	Component rule using
	Classification		Classification (concept learning) Instance Classification	D Set and/or R Set Classification	R Sets and/or D Sets and/or Operations Classification
	Discriminated Recall	Discrete Memory (Discrimination Learning) Component Recall	Definition and/or Instance Recall	Rule and/or Solution Recall	Higher Rule and/or Solution Recall

Content Dimension

*Notes:

1. The diagonal (shaded boxes) are the categories included in the previous Gagné-Merrill hierarchy (Gagné 1965, 1970; Merrill 1971). Parentheses are Gagné's terms. Italicized words are suggested new terms used in this paper to classify tasks in both dimensions simultaneously.

Footnotes

[1]A discussion of attribute matching appears later in the paper.

[2]Măth thēm a′gĕn ĭc was coined by Ernest Rothkopf (1965). The roots of this word are mathema: learning, and gignesthai: to be born. Mathemagenic seems appropriate as an adjective which describes behaviors which give birth to learning or information which is designed to promote those behaviors. The use of the word here is an extension of Rothfopf's original application but it seems consistent with his intent.

[3]Algorithms are defined as a set of procedures usually involving a repetition of some operation for solving a mathematical problem. The word has been generalized in this scheme to refer to procedures for solving any kind of problem and includes both production algorithms and search algorithms.

[4]While some symbols have been used for presentation form and others suggested for illustrative purposes here, there has not been an attempt to suggest or promote a set of symbol conventions for the variables suggested. Such a system will await further development of this taxonomy.

References

Anderson, R. C. How to construct achievement tests to assess comprehension. Review of Educational Research, 1972, in press. (a)

Anderson, R. C. Learning principles from text. Journal of Educational Psychology, 1972, in press. (b)

Anderson, R. C., & Kulhavy, R. W. Learning concepts from definitions. American Educational Research Journal, 1972, in press.

Baker, E. L. Effects of student achievement on behavioral objectives and non-behavioral objectives. Journal of Experimental Education, 1969, 37, 5-8.

Bloom, B. S. (Ed.) Taxonomy of educational objectives: The classification of educational goals. Handbook 1. Cognitive domain. New York: McKay, 1956.

Bower, G. H. Analysis of a mnemonic device. American Scientist, LVIII (September-October, 1970) 496-510.

Briggs, L. J. Sequencing of instruction in relation to hierarchies of competence. Pittsburgh: American Institutes for Research, Monograph No. 3, 1968.

Briggs, L. J. Handbook of procedures for the design of instruction. Pittsburgh: American Institutes for Research, Monograph No. 4, 1970.

Clark, D. C. Teaching concepts in the classroom: A set of teaching prescriptions derived from experimental research. Journal of Educational Psychology Monographs, 1971, 62, 253-278.

Cook, M. J. Behavioral objectives and rate of forgetting. Paper presented at the meeting of the American Educational Research Association, Baltimore, Baltimore County, March 1970.

Craig, R. C. Directed versus independent discovery of established relations. Journal of Educational Psychology, 1956, 47, 223-234.

Dalis, G. T. Effect of precise objectives upon student achievement on health education. Journal of Experimental Education, 1970, 39, 20-23.

Frase, L. T. Some unpredicted effects of different questions upon learning from connected discourse. Journal of Educational Psychology, 1968 59, 197-201. (a)

Frase, L. T. Effect of question location, pacing, and mode upon retention of prose material. Journal of Educational Psychology, 1968, 59 244-249. (b)

Gagné, R. M. The conditions of learning. New York: Holt, Rinehart and Winston, 1965.

Gagné, R. M. The conditions of learning. (2nd ed.) New York: Holt, Rinehart and Winston, 1970.

Gagné, R. M., & Brown, L. T. Some factors in the programming of conceptual learning. Journal of Experimental Psychology, 1961, 62, 313-321.

Gagné, R. M., & Rohwer, W. D., Jr. Instructional psychology. Annual Review of Psychology, 1969, 20, 381-418.

Gerlack, V. S., & Ely, D. D. Teaching and media: A systematic approach. Englewood Cliffs: Prentice-Hall, 1971.

Glaser, R. Toward a behavioral science base for instructional design. In R. Glaser (Ed.), Teaching machines & programmed learning, II, Washington, D. C.: National Education Association, 1965.

Glaser, R. The design of instruction. The Changing American School: NSSE 65th Yearbook, Chicago: University of Chicago Press, 1966.

Glaser, R. Instructional technology and the measurement of learning outcomes: Some questions. In W. J. Popham (Ed.), Criterion-referenced measurement: An introduction. Englewood Cliffs: Educational Technology Publications, 1971.

Glaser, R., & Klaus, D. J. Proficiency measurement: Assessing human performance. In R. Gagné (Ed.), Psychological principles in system development, New York: Holt, Rinehart and Winston, 1962.

Glaser, R., & Nitko, A. J. Measurement in learning and instruction. In R. L. Thorndike (Ed.), Educational measurement, Washington, D. C.: American Council Education, 1971, 625-670.

Glaser, R., & Resnick, L. B. Instructional psychology, In P. H. Mussen & M. R. Rosenzweig (Eds.), Annual Review of Psychology, 1972, 23.

Guthrie, J. T. Expository instruction versus a discovery method. Journal of Educational Psychology, 1967, 58, 45-49.

Guthrie, J. T., & Baldwin, T. L. Effects of discrimination, grammatical rules, and application of rules on the acquisition of grammatical concept.

Journal of Educational Psychology, 1970, 61, 358-364.

Haslerud, G. M., & Meyers, S. The transfer value of given and individually derived principles. Journal of Educational Psychology, 1958, 49, 293-298.

Hiscox, M. D., & Herron, D. J. The ability of students to correlate specific and nonspecific educational objectives with test questions. Unpublished manuscript, Purdue University, 1971.

Horton, G. A., Jr. Spanish language program development, validation and subtractive testing of a learning set. (Doctoral dissertation, University of California at Los Angeles), 1969.

Jenkins, J. R., & Deno, S. L. Influence of knowledge and type of objectives on subject matter learning. Journal of Educational Psychology, 1971, 62, 67-70.

Jenkins, J. R., & Niesworth, J. T. The facilitative influence of instructional objectives. Journal of Educational Research, in press.

Johnson, D. M., & Stratton, R. P. Evaluation of five methods of teaching concepts. Journal of Educational Psychology, 1966, 57, 48-53.

Kemp, J. E. Instructional design: A plan for unit and course development. Belmont, California: Fearon, 1971.

Kersh, B. Y. The adequacy of "meaning" as an explanation for the superiority of learning by independent discovery. Journal of Educational Psychology, 1958, 49, 282-292.

Kersh, B. Y. The motivating effect of learning by directed discovery. Journal of Educational Psychology, 1962, 53, 65-71.

Kulhavy, R. W., & Anderson, R. C. The delay-retention effect with multiple-choice tests. Journal of Educational Psychology, 1972, in press.

Mager, R. F., & McCann, J. Learner controlled instruction. Mimeographed paper, Palo Alto: Varian Associates, 1961, iii-21.

Markle, S. M., & Tiemann, P. W. Some principles of instructional design at higher cognitive levels. In B. Rollett and K. Weltner (Ed.), Fortschritte und ergebnisse der unterrichtstechnologie, Berlin; Ehrenwirth-Verlag, 1972, in press.

McNeil, J. D. Concomitants of using behavioral objectives on the assessment of teacher effectiveness. Journal of Experimental Education, 1967, 36, 69-74.

Merrill, M. D. Correction and review on successive parts in learning a hierarchical task. Journal of Educational Psychology, 1965, 56, 225-234.

Merrill, M. D. Specific review versus repeated presentation in a programmed imaginary science. Journal of Educational Psychology, 1970, 61, 392-399.

Merrill, M. D. Instructional design: Readings, Englewood Cliffs, N. J.: Prentice-Hall, 1971. (a)

Merrill, M. D. Necessary psychological conditions for defining instructional objectives. Educational Technology, 1971, 2 (8), 34-40. (b)

Merrill, M. D. Toward a theory-based approach to instructional development. Working Paper No. 16, Brigham Young University, 1971. (c)

Merrill, M. D. Psychomotor taxonomies, classification, and instructional theory. In R. Singer (Ed.), The psychomotor domain movement behavior, Chapter 14, Philadelphia: Lea and Febiger, 1972, in press.

Merrill, M. D., & Stolurow, L. M. Hierarchical preview versus problem-oriented reivew in learning an imaginary science. American Educational Research Journal, 1966, 3, 251-261.

Merrill, M. D., & Tennyson, R. D. Concept acquisition and specified errors as a function of relationships between positive and negative instances. American Educational Research Journal, 1972, in press. (a)

Merrill, M. D., & Tennyson, R. D. Attribute prompting variables in learning classroom concepts. Journal of Educational Psychology, 1972, in press. (b)

Merrill, M. D., Barton, K., & Wood, L. E. Specific review in learning a hierarchical imaginary science. Journal of Educational Psychology, 1970, 61, 102-109.

Merrill, P. F. The effects of the availability of objectives and/or rules on the learning process. Paper presented at the meeting of the American Educational Research Association, New York, February, 1971.

Osler, S. F., & Fivel, M. W. Concept attainment: I. The role of age and intelligence in concept attainment by induction. Journal of Experimental Psychology, 1961, 62, 1-8.

Oswald, J. M. Some effects of instructor specified instructional objectives on achievement of knowledge and comprehension. Dissertation Abstracts International, 31 (8), 1971, Order No. 71-2865.

Piatt, R. G. An investigation of the effects on learning outcomes of the training of teachers on defining, writing, and implementing behavioral objectives. Educational Technology Research, No. 8, Englewood Cliffs, Educational Technology Publications, 1971.

Popham, J. W., & Baker, E. L. Systematic instruction, Englewood Cliffs: Prentice-Hall, 1970.

Roughead, W. G., & Scandura, J. M. 'What is learned' in mathematical discovery. Journal of Educational Psychology, 1968, 59, 283-289.

Rohwer, W. D., Jr. Individual differences. Paper presented at the meeting of the American Educational Research Association, Minneapolis, Minnesota, 1970.

Scandura, J. M. Abstract and tasks for use in problem-solving research. The Journal of Experimental Education, 1964, 33, 145-148.

Scandura, J. M. Algorithm learning and problem-solving. The Journal of Experimental Education, 1966, 34, 1-6. (a)

Scandura, J. M. Prior learning, presentation order, and prerequisite practice in problem-solving. The Journal of Experimental Education, 1966, 34, 12-18. (b)

Scandura, J. M. Precision in research on mathematics learning: The emerging field of psycho-mathematics. Journal of Research in Science Teaching, 1966, 4, 253-274. (c)

Scandura, J. M. New directions for theory and research on rule learning 1. A set function language. Acta Psychologica, 1968, 28, 301-321.

Scandura, J. M. Role of rules in behavior toward an operational definition of what (rule) is learned. Psychological Review, 1970, 77, 516-533.

Scandura, J. M., & Dunnin, J. H. Extra-scope transfer in learning mathematical strategies. Journal of Educational Psychology, 1968, 59, 350-354.

Scandura, J. M., & Voorhies, D. J. Effects of irrelevant attributes and irrelevant operations on rule learning. Journal of Educational Psychology, 1971, 62, 352-356.

Scandura, J. M., Woodward, E., & Lee, F. Rule generality and consistency in mathematics learning. American Educational Research Journal, 1967, 4, 303-319.

Sudweeks, R., & Merrill, M. D. The effects of component concepts classification errors in subsequent learning of a defined concept. Journal of Educational Psychology, submitted, 1972.

Tennyson, R. D. Correct classification, overgeneralization, undergeneralization, and misconception as a function of probability, divergence, and matching of instances in concept acquisition. Journal of Educational Psychology, submitted, 1972.

Tennyson, R. D., & Boutwell, R. C. Instance probability analysis, Harvard Educational Review, submitted, 1972.

Tennyson, R. D., & Woolley, F. R. Conceptual model of classification behavior. Educational Technology, 1972, in press.

Tennyson, R. D., Woolley, F. R., & Merrill, M. D. Exemplar and non-exemplar variables which produce concept classification behavior and specified classification errors. Journal of Educational Psychology, 1972, 63 (2), 144-152.

Thatcher, D. L. The use of examples and nonexamples in principle learning behavior. Unpublished manuscript, Brigham Young University, 1972.

Traub, R. E. Importance of problem heterogeneity to programmed instruction. Journal of Educational Psychology, 1966, 57, 54-60.

Wang, M. C., Resnick, L. B., & Boozer, R. The sequence of development of some early mathematics behaviors. Child Development, June, 1972.

Watts, G. H., & Anderson, R. C. Effects of three types of inserted questions on learning from prose. Journal of Educational Psychology, 1971, 62, 387-394.

Wittrock, M. C. Verbal stimuli in concept formation: Learning by discovery. Journal of Educational Psychology, 1963, 54, 183-190.

Woolley, F. R. Effects of the presence of concept definition, pretraining, concept exemplars, and feedback on the instructor of infinite conjunction concepts. (Doctoral dissertation, Brigham Young University) Provo, Utah, 1971, 177 pp.

Worthen, B. Discovery and expository task presentation in elementary mathematics. Journal of Educational Psychology: Monograph Supplement, 1968, 59, 1-13.

Young, J. I., Smith, K. H., & Merrill, M. D. The effects of review techniques and instance presentation on concept learning tasks. (Doctoral dissertation, Brigham Young University) Provo, Utah, 1972.